

# Prédiction du prix de l'or

## Le contexte

Historiquement, l'or a été utilisé comme une forme de monnaie dans diverses parties du monde, y compris aux États-Unis. À l'heure actuelle, les métaux précieux comme l'or sont détenus auprès des banques centrales de tous les pays pour garantir le remboursement des dettes étrangères et également pour contrôler l'inflation, ce qui reflète la solidité financière du pays. Récemment, les économies mondiales émergentes, telles que la Chine, la Russie et l'Inde ont été de gros acheteurs d'or, tandis que les États-Unis, l'Afrique du Sud et l'Australie sont parmi les gros vendeurs d'or.

Les prévisions de hausse et de baisse des taux quotidiens de l'or peuvent aider les investisseurs à décider quand acheter (ou vendre) la matière première. Mais les prix de l'or dépendent de nombreux facteurs tels que les prix d'autres métaux précieux, les prix du pétrole brut, les performances boursières, les prix des obligations, les taux de change, etc.

Le défi de ce projet est de prédire avec précision le futur cours de clôture ajusté de l'ETF sur l'or sur une période donnée dans le futur.

## Présentation des données

Les données de cette étude sont recueillies du 18 novembre 2011 au 1er janvier 2019 à partir de diverses sources. Les données ont 1718 lignes au total et 80 colonnes au total. Données pour les attributs, tels que le prix du pétrole, l'indice Standard and Poor's (S&P) 500, les taux des obligations américaines de l'indice Dow Jones (10 ans), les taux de change Euro USD, les prix des métaux précieux argent et platine et d'autres métaux tels que le palladium et le rhodium, Les prix de l'indice du dollar américain, de l'Eldorado Gold Corporation et de l'ETF Gold Miners ont été rassemblés.

L'ensemble de données comporte 1718 lignes au total et 80 colonnes au total. Données pour les attributs, tels que le prix du pétrole, l'indice Standard and Poor's (S&P) 500, les taux des obligations américaines de l'indice Dow Jones (10 ans), les taux de change Euro USD, les prix des métaux précieux argent et platine et d'autres métaux tels que le palladium et le rhodium, Les prix de l'indice du dollar américain, de l'Eldorado Gold Corporation et de l'ETF Gold Miners ont été rassemblés.

Les données historiques de Gold ETF extraites de Yahoo finance ont 7 colonnes, Date, Open, High, Low, Close, Adjusted Close et Volume, la différence entre Adjusted Close et Close est que le prix de clôture d'une action est le prix de cette action à la clôture du jour de bourse. Alors que le cours de clôture ajusté prend en compte des facteurs tels que les dividendes, les fractionnements d'actions et les nouvelles offres d'actions pour déterminer une valeur. Nous utiliserions la clôture ajustée comme variable de résultat, ce qui correspond à la valeur que nous voulons prédire.

'SPopen', 'SPhigh', 'SPlow', 'SPclose', 'SPAjclose', 'SPvolume' de l'indice S&P 500,

'DJopen', 'DJhigh', 'DJlow', 'DJclose', 'DJAjclose', 'DJvolume' de l'indice Dow Jones,

'EGopen', 'EGhigh', 'EGlow', 'EGclose', 'EGAjclose', 'EGvolume' d'Eldorado Gold Corporation (EGO),

'EUPrice', 'EUOpen', 'EUHigh', 'EULow', 'EUTrend' du taux de change EUR USD,

'OFPrice', 'OFOpen', 'OFHigh', 'OFLow', 'OFVolume', 'OFTrend' des contrats à terme sur le pétrole brut Brent,

'OSPrice', 'OSOpen', 'OSHigh', 'OSLow', 'OSTrend', du pétrole brut WTI USD,

'SFPrice', 'SFOpen', 'SFHigh', 'SFlow', 'SFVolume', 'SFTrend' de Silver Futures,

'USBPrice', 'USBOpen', 'USBHigh', 'USBLow', 'USBTrend' des données de taux des obligations américaines,

'PLTPrice', 'PLTOpen', 'PLTHigh', 'PLTLow', 'PLTTrend' du prix du platine,

'PLDPrice', 'PLDOpen', 'PLDHigh', 'PLDLow', 'PLDTrend' du prix du palladium 'RHOPRICE' des prix du rhodium

'USDIPrice', 'USDIOpen', 'USDIHigh', 'USDILow', 'USDIVolume', 'USDITrend' du prix de l'indice du dollar américain,

'GDXOpen', 'GDXHigh', 'GDXLow', 'GDXClose', 'GDXAdj Close', 'GDXVolume' de l'ETF Gold Miners,

'USOOpen', 'USOHigh', 'USOLow', 'USOClose', 'USOAdj Close', 'USOVOLUME' de l'ETF pétrolier USO sont utilisés comme variables d'ingénierie de caractéristiques.

## Consignes

1. Faire l'analyse exploratoire du Dataset
  - a. Identification de la target
  - b. Nombre des lignes et de colonnes
  - c. Identification des valeurs manquantes
  - d. Types de variables
  - e. Visualisation de la cible (target)
  - f. Compréhension des différentes variables
  - g. Visualisation des relations : features/target
  - h. Identification des outliers
2. Prédire le prix de l'or
  - a. Pre-processing
  - b. Modélisation